REMARKS

The fact that July 9, 2006, fell on a Sunday ensures that this paper is timely filed as of Monday, July 10, 2006, the next business day. Applicants and the undersigned are most grateful for the time and effort accorded the instant application by the Examiner. The Office is respectfully requested to reconsider the rejection present in the outstanding Office Action in light of the following remarks.

In the Office Action dated February 9, 2006, pending Claims 1-19 were rejected and the rejection made final. In response Applicants filed an Amendment After Final and received an Advisory Action. Applicants' counsel also conducted an interview with the Examiner on July 10, 2006. No agreement was reached with respect to the claims. Applicants are herewith filing a Request for Continued Examination (RCE) and respectfully request the Office to reconsider the rejections presented in the outstanding Office Action in light of the following remarks.

As a preliminary matter, Applicants note the Office has not yet acknowledged the claim of priority in this case, not the submission of the certified copy of the priority document. Such acknowledgement is respectfully requested in the next communication from the Office.

Claims 1-19 were pending in the instant application at the time of the outstanding Office Action. Of these claims, Claims 1, 10, and 19 are independent claims; the remaining claims are dependent claims. Claims 1, 10, and 19 have been rewritten. Applicants intend no change in the scope of the claims by the changes made by these

Atty. Docket No. JP920000191US1
(590.079)

amendments. It should also be noted these amendments are not in acquiescence of the Office's position on allowability of the claims, but merely to expedite prosecution.

Claims 1-3, 6-12, and 15-19 stand rejected under 35 USC § 103(a) as being unpatentable over Wang et al. (hereinafter "Wang") in view of Razin et al. (hereinafter "Razin"). Reconsideration and withdrawal of the present rejections are hereby respectfully requested.

The present invention is directed to a method and apparatus for automatically extracting new words from a cleaned corpus, where the corpus can be in any language that may or not have word boundaries (ranging from English or Latin to Chinese or Japanese). The instant invention segments a cleaned corpus to form a segmented corpus, splits the segmented corpus to form sub strings, and counts the occurrences of each sub strings appearing in the given corpus. Finally, the present invention filters out false candidates to output new words.

As best understood, Wang appears to be directed to a method that optimizes language models in which an initial language model is developed from a lexicon and segmentation derived from a received corpus. The initial model is iteratively refined by updating the lexicon and re-segmenting the corpus using both maximum match techniques and statistical principles. (Abstract) As asserted in the outstanding Office Action, Wang does not expressly disclose filtering out false candidates to output new words. Further, Wang does not expressly disclose that the segmenting and the splitting of the corpus is not dependent upon word boundaries.

- 8 -

Razin fails to overcome the deficiencies of Wang as set forth above. As best understood, Razin appears to be directed to standardizing phrasing in a document. Razin identifies phrases in a document to create a preliminary list of phrases, then filters and refines those phrases to create a final list of standard phrases. Razin then identifies phrase of a document that are similar to standard phrases, decides if the candidate phrase is similar enough to the standard phrase and compute phrase substitutions to determine the approximate conformation of the standard phrase to the approximate phrase and vice versa. (Abstract) There is no suggestion or teaching in Razin that the segmenting and the splitting of the corpus is not dependent upon word boundaries. In fact, Razin teaches away from this ability (column 11, lines 14-36), teaching that the source text is segmented using a standard finite-state machine technique that recognizes patterns that indicate word and sentence boundaries.

Claim 1 recites a "method of extracting new word automatically, said method comprising the steps of: segmenting a cleaned corpus to form a segmented corpus; splitting the segmented corpus to form sub strings, and counting the occurrences of each sub strings appearing in the corpus; and filtering out false candidates to output new words; **wherein the segmenting and the splitting is not dependent upon word boundaries.** (emphasis added) Similar language also appears in the other Independent Claims. Neither Wang nor Razin, nor the combination of the two, teach or suggest the limitations of the instant invention.

Further, a 35 USC 103(a) rejection requires that the combined cited references provide both the motivation to combine the references and an expectation of success. Not

only is there no motivation to combine the references, no expectation of success, but actually combining the references would not produce the claimed invention. Thus, the claimed invention is patentable over the combined references and the state of the art.

Claims 4-5 and 13-14 stand rejected under 35 USC § 103(a) as being unpatentable over Wang et al. (hereinafter "Wang") in view of Razin et al. (hereinafter "Razin") and further in view of Hui. Specifically the Office asserted that "[i]t would have been obvious ... to modify Wang in view of Razin by specifically providing using extended suffix tree (GST or GAST), for the purpose of storing more than one input strings." Reconsideration and withdrawal of this rejection is hereby respectfully requested.

Hui does not overcome the deficiencies of Wang or Razin. As best understood, Hui is directed towards an algorithm that provides an optimal sequential solution of the color set size problem which entails finding the number of different leaf colors in a subtree rooted at a vertex v in a rooted tree. Although Hui asserts that there is applicability in string matching heuristics, there is no teaching or suggestion in Hui that the segmenting and the splitting of the corpus is not dependent upon word boundaries.

Combining Wang, Razin, and Hui would result in producing a language model of phrases using an optimal sequential solution to find the phrases that constitute the lexicon of standard phrases. Even if there were a motivation for the combination, this combination does not teach or suggest the claimed invention.
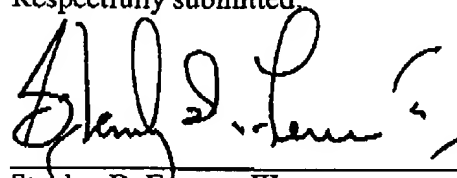
In view of the foregoing, it is respectfully submitted that Independent Claims 1, 10 and 19 fully distinguish over the applied art and are thus allowable. By virtue of

- 10 -

Atty. Docket No. JP920000191US1
(590.079)

dependence from Claims 1 and 10, it is thus also submitted that Claims 2-9 and 11-18 are also allowable at this juncture.

In summary, it is respectfully submitted that the instant application, including Claims 1-19, is presently in condition for allowance. Notice to the effect is hereby earnestly solicited. If there are any further issues in this application, the Examiner is invited to contact the undersigned at the telephone number listed below.

Respectfully submitted,

Stanley D. Ference III
Registration No. 33,879

Customer No. 35195
FERENCE & ASSOCIATES
409 Broad Street
Pittsburgh, Pennsylvania 15143
(412) 741-8400
(412) 741-9292 – Facsimile

Attorneys for Applicants

- 11 -